

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: July 18, 2001, 15:56:47 ; Search time 19.26 Seconds

(without alignments)
3021.670 Million cell updates/sec

Title: US-09-587-111-5

Perfect score: 4004
Sequence: 1 MTPSSSPVFLRLETLDDGQE.....EDEDGASENVVPLVOLLQSN 764Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 200000000Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1652	41.3	838	2 T09054	capsaicin receptor
2	634	15.8	725	3 JC7531	calcium transport
3	414	10.3	900	2 T33026	hypothetical prote
4	392.5	9.8	790	2 T20312	hypothetical prote
5	364	9.1	937	2 T37241	olfactory channel
6	261	6.5	519	2 T24772	hypothetical prote
7	209	5.2	1188	2 T19552	hypothetical prote
8	197.5	4.9	1124	2 JH0588	calmodulin-binding
9	191	4.8	1274	2 JN0015	trp protein - frui
10	190.5	4.8	1275	2 JU0092	trp protein - frui
11	174	4.3	810	2 T38361	TRPC1 protein - hu
12	167.5	4.2	934	1 H71274	probable ankyrin -
13	157.5	3.9	481	2 T23729	hypothetical prote
14	157	3.9	1549	2 T13940	ankyrin - fruit fl
15	151	3.8	1765	2 T42714	ankyrin 3, splice
16	151	3.8	1940	2 T42715	ankyrin 3, splice
17	151	3.8	1943	2 T42713	ankyrin 3, splice
18	151	3.8	1961	2 T42716	ankyrin 3, splice
19	151	3.8	3924	2 S37431	ankyrin 2, neurona
20	150.5	3.8	793	2 S68338	trp-1 protein - hu
21	150	3.7	887	2 T03939	potassium channel
22	150	3.7	4377	2 A55575	ankyrin 3, long sp
23	147	3.7	1001	2 S30385	ankyrin 3, splice
24	143	3.6	842	2 T32258	G9a protein - huma
25	141.5	3.5	683	2 A85044	hypothetical prote
26	140	3.5	1848	2 S37771	hypothetical prote
27	140	3.5	1862	2 T49502	ankyrin - erythrocy
28	139.5	3.5	1411	2 S30355	ankyrin - mouse
29	138	3.4	1856	2 B35049	alpha-latroinsecto ankyrin 1, erythro

30	138	3.4	1880	2 A35049	ankyrin 1, erythro
31	138	3.4	1881	1 SJHUK	ankyrin 1, erythro
32	136	3.4	2352	2 T30201	Notch homolog prot
33	135	3.4	828	2 JC5807	trp3 protein - rat
34	134.5	3.4	1401	2 S11527	alpha-latrotoxin p
35	134.5	3.4	1435	2 T32930	hypothetical prote
36	133.5	3.3	642	2 S58154	hypothetical prote
37	133.5	3.3	1786	2 A57282	ankyrin-related pr
38	133.5	3.3	1815	2 T15346	elegans ankyrin-re
39	133.5	3.3	1867	2 T15344	ankyrin-related un
40	133.5	3.3	2039	2 T15347	ankyrin-related un
41	132	3.3	616	2 T00884	hypothetical prote
42	132	3.3	1184	2 T00253	gene Anknzn protei
43	132	3.3	1398	2 T21884	hypothetical prote
44	132	3.3	1964	2 T09059	notch4 - mouse
45	131.5	3.3	598	2 D71127	hypothetical prote

ALIGNMENTS

RESULT 1

capsaicin receptor - rat
N:Alternate names: vanilloid receptor subtype 1
C:Species: Rattus norvegicus (Norway rat)
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 21-Jul-2000
C:Accession: T09054
R:Catterina, M.J.; Schumacher, M.A.; Tomlinage, M.; Rosen, T.A.; Levine, J.D.; Julius, Nature 389, 816-824, 1997
A:Title: The capsaicin receptor: A heat-activated ion channel in the pain pathway.
A:Reference number: Z16539; MUID:98007969
A:Accession: T09054
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1838 <CAT>
A:Cross-references: EMBL:AF029310; MID:92570932; PIDN:AAC53398.1; PID:92570933
A:Experimental source: dorsal root ganglion
C:Keywords: ion channel; receptor

Query Match	41.3%	Score 1652;	DB 2;	Length 838;
Best Local Similarity	46.3%	Pred. No. 4.4e-120;		
Matches 359;	Conservative 127;	Mismatches 230;	Indels 60;	Gaps 15;
QY	18	GOEDGSEADRGKIDFGSGLPMEISQFQEDKKRFAQIRVN---	LNRYKRGTC-----	65
DB	51	GKGDSEERS-----	PLDCPYEKGGLASCPITIVSSVLTIOKRGDPAASVPSQ	99
QY	66	---ASQPDNRFDRDLFNFAVSGVPEDLAFLPEYLSKTSKYLTDSYETGSGTKCLM		121
DB	100	DSVAGAKRPRLRYRSTFDFVAOSNCOELESLLPFIQSRSKRLTDSFNPREGKTCIL		159
QY	122	KAVLNLDGVNACILPLLDIDRDSGNPQPLVNAOCTDDYRGRSALHAIKRSLOCVKL		181
DB	160	KAMLNHNGQNDITALLLDVARKTDSLKQFVNAAGYTFSGQFALHAIERRNNLTVTLL		219
QY	182	LVENGAVVHARACGRFQKGG--TCFYEGELPLSLACTOMDVSVLLEMPHOPASLOA		240
DB	220	LVENGADVQAANAGDFRFRKTRGPGFYFEGELPLSLACTOALAVKLLQNSWCPADISA		279
QY	241	TDSQGNVLAHLVWISNSAENIALVTSMYGDLQAGARLCPTVQLDRIKLDITPLKL		300
DB	280	RDSVGNVLAHLVAVANTYDNTKFTVSMVNEIILLAKLHPTLKEITIRKCLTFLAL		339
QY	301	AAKGGKIEIRHIIQREP--SGLSLSRKFTKWCYGVRSVLDLASVDSCEANSVLEII		358
DB	340	AASGKIGVLAIIYIQRREIHEPECRLSRKFTFMAVGVHSLYDLSCIDCEKNSVLEVI		399
QY	359	AF-HCKSPHRRBMVLEPLKLLQAKMDLLPK--FELNPLCNLYMFLPAVAVHOPTLK		416
DB	400	AYSSSETPNRHMDLVEPLNKLQDKMDRVPVKRIEYFNVFYCYLYMIIITFAAAAYRPV--		457

QY	417	KQAAI-HKKAEVSGSMIIIGIIILLGGIYLLVAGQIAEMRRHHVFIWISFDSEFFL	474
Db	458	-EGLPPLKNTKVGDIERYKTEIISVSGGVYFFERIGDIYFLQRPDSLSEVDSSEFL	516
QY	475	LEQALLFVVSQVLCFLAIEWYLPILVSAIVLGWLNLLYYRGFOHTGIYSVMIQVILRD	534
Db	517	FVQSEFMFLSVLVYRSQREKVEYASWVFLSAGWIMLTLYTGFOOMGIYAVIMEIRLD	576
QY	535	LLRFLILYLVFLFGAVVALVSLSOEAMRPEAPGTGNATNESVQPMQGDSENGA-----	588
Db	577	LCRFHEVVLVEFGSTAVVILIED-----GKN---NSLPHESTPHKRCGSAACKGN	625
QY	589	QYRGILEASLELEKFTTIGMELAFQEOHFRGMVLLLLLAVVLTYYILLNMLALMSET	648
Db	626	SYNSLYSTCLELFEKFTTIGMDLEFENYDFKAVFTILLLAVVILTYILLNMLALMSET	685
QY	649	VNSVATDSMSIMIKLOKALSVLEMENGYMMC-RKKORACVMLTVGKPPGSDPERCPEVE	707
Db	666	VNKIKQESKNIKLQORALTIIDLTESFLKCRMKAFRRSKILLQVGTTPGCKDDYRKCPEVD	745
QY	708	EYVNASMEQTLPLEICEDPSGA-GVPRILENPVLASPKPEDEGDGASENYVPVQLIQ	762
Db	746	EYVNTTNTNTNVGQIINEDPGNCEGVKRTISFLSRG----RVSGRMKKMFALVPLLR	797

RESULT 2
JC7531
calcium transport protein, Cat1 - human
C:Species: Homo sapiens (man)
C:Date: 31-Mar-2001 #sequence-revision 31-Mar-2001 #text-change 31-Mar-2001
C:Accession: JC7531
R:Peng, J.B.; Chen, X.; Berger, U.V.; Weremowicz, S.; Morton, C.C.; Vassiliev, P.M.; Brown, B.
Biochem. Biophys. Res. Commun. 278, 326-332, 2000
A:Title: Human calcium transport protein Cat1.
A:Reference number: JC7531; MUID:20551480
A:Accession: JC7531
A:Molecule type: mRNA
A:Residues: 1-725 <PDB>
A:Cross-references: GB:AF304463
C:Comment: This protein, a member of a family of Ca2+ channels, has a role in cellular
line and kidney.
C:Genetics:
A:Gene: Cat1
A:Map position: 7q33-34
A:Keywords: calcium channel; calcium transport; intestine; kidney; transport protein; trn;

[illegible]

```

Db      331  MGLAIYLLIICFLMCCITARYPLKPRNTNNKSRPNDNTLLQOKLLQEAHYMMPRKDDIKUYEL 390
OY      438  LILGGIYLLVGLQOLMYWRRHV--FWIWSIFDSTFEIILFQALITVVSQVLCFLAIEMY 495
::      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      391  VTVIGAIITILVEVPDIFERRGVTRFFEQOTLLGPRFHVLIITAFVWLIVMYRRLISASGE 450
OY      496  LPLVLSALVIGMLNLITVYTRGFQHTGYSVMIOKVLRLDRLRFPLLYLFLFEGFVALVS 555
::      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      451  VYPSFALVYLQMCVMYFANGFQMLGFFTIIDMKILFGDLMRCHWLMAVILIGFSAFYI 510
OY      556  LSGEARREPAATGPNATESVQPMGQEDBSGNAQYKGLLEASILELFPKTLTGCELAFQEQ 615
::      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      511  IFQ-----TEDPEELGHFYDYPMALEFSTFELR-LIIIDGPRANYVD 550
OY      616  LHFPGVYLLLLAVLLTYLITLLNMLIALMSETVNSVAVDMSIMYLQKAIIVLEMEY 675
::      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      551  LPF--MSYITAAAIITITLLMLNLLLALMMGDTIHWKVAHREDEIMQAQIVATTVIMIERL 608
OY      676  ---WM-----CRKKORAGVMLVGTKPDGSDPERMCFRYDE 708
::      ::      ::      ::      ::      ::      ::      ::      ::      ::      ::
Db      609  PRCLMPRSGIGRGYEGG-----DRLWLLRYED 635

```

RESULT 3
 T33026
 hypothetical protein T09A12.3 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T33026
 R:Hawkins, J.; Fulton, B.; Gillam, B.
 submitted to the EMBL Data Library, February 1998
 A:Description: The sequence of *C. elegans* cosmid T09A12.
 A:Reference number: 221265
 A:Accession: T33026
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-900 <HAM>
 A:Cross-references: EMBL:AF047660; PIDD:AA004431.1; GSPDB:GN00022; CESP:T09A12.3
 A:Experimental source: strain Bristol N2; clone T09A12
 C:Genetics:
 A:Gene: CESP:T09A12.3
 A:Map position: 4
 A:Introns: 43/2; 86/3; 260/2; 396/3; 424/2; 495/1; 517/3; 553/3; 691/3; 741/3; 780/3;

[illegible]

QY 376 LNLQAKMILLIR-FFLNLCLMLITMFTTAV-----AYHOP-- 413
 Db 409 LDTLLEAKWEFAKRNMIISFTALTYICVTAFTLPIGFSTEMLTGWINRYSEBFP 468
 QY 414 -----FLKKAAP-----HKAKEVGSMLTGIIILLGIGIYLLV 448
 Db 469 GRYKNSSTLQOVKVINATSGVLWSEPPISQCHLRNWDPDIPRANSYIRLVELEFVI 528
 QY 449 GOLYFMRHFWIWFID-----SYFELLFQALLT-----VSOVLCELAIEW 494
 Db 529 GIC-----VQVFLDFRDIKIRGKKMMNVLTAFPAKITFKLYFLVLAIMPRLACDL 581
 QY 495 YLPLLY-----SALVGLMNLTYTRGEQHTGISVMIOKVILRDLRLIYVLF 547
 Db 582 SPVLVVDNVLITVMTFTVHYLYCVRVIRFVGPFLMYTIIATDIFRMLIYGFILM 641
 QY 548 GF--AVLVLSLOEAMRPEATPTGNATESVOPMEGOEDE-----GNGQX 590
 Db 642 GFSGFSLIFLSCF--RANAVIKKLITDOSEASGSDKNKMLTROISAVDTAIVKMAEVF 699
 QY 591 RGLDEASLELF--KFTIGMELAFQEOLEHFRGMVL-----LILLAYVLLTYILL 637
 Db 700 ENVMQSPLEAVRFTFILLIGEFT----VLYRNALALCPANFMWIGKVFIIEFLFVSLMQ 755
 QY 638 LNMILALMSETVNSVADTSMISIKLQKAIISYLEME 672
 Db 756 FNNLIAMTTRTYETIFQTOLE-YKRORAOVILME 789

RESULT 4

T20312
 hypothetical protein F28H7.10 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T20312; T21533
 R:Matthews, P.
 Submitted to the EMBL Data Library, June 1996
 A:Accession number: Z19255
 A:Accession: T20312
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-790 <WIL>
 A:Cross-references: EMBL:Z74030; PIDN:CAA98449.1; GSPDB:GN00023; CESP:F28H7.10
 A:Experimental source: clone D1054
 R:Berts, M.
 Submitted to the EMBL Data Library, May 1996
 A:Reference number: Z19435
 A:Accession: T21533
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-790 <W12>
 A:Cross-references: EMBL:Z72508; PIDN:CAA96644.1; GSPDB:GN00023; CESP:F28H7.10
 A:Experimental source: clone F28H7
 C:Genetics:
 A:Gene: CESP:F28H7.10
 A:Map position: 5
 A:Introns: 46/2; 89/3; 129/1; 157/1; 201/1; 264/2; 349/3; 406/3; 487/3; 543/2; 580/1; 67

Query Match 9.8%; Score 392.5; DB 2; Length 790;
 Best Local Similarity 25.3%; Pred. No. 2.6e-22;
 Matches 182; Conservative 106; Mismatches 243; Indels 189; Gaps 35;
 QY 61 RKGASGSPD-PNRFDDRLFNNAVSRGPEDLAIGLPELSTSKYLL-----TDSYXT 111
 Db 101 KKGKSGKPNLIDFDDGO-----AEMAG--DLKALKLLDGGGKGSSEKTYR 147
 QY 112 E-----GSTGKT-----CLMK--AVLNLDGVNACTLPLQIDRDSGNQP--LVNAQC 156
 Db 148 EWMVSVDERGSMENLAILCIGGSAHLNL-----IARLRINFPLINDIC 194
 QY 157 TDDYIRGHSAIHAIERKSLQCVKLIVENGANVHARACGRFF-----OKGCT----- 204

Db 195 VSEYYGLSPLHLAIVNDOAFTSLRLGLADLNORCYGAFPCADDOKASTDSEHEFV 254
 QY 205 -----CFYFGLPELPLSLACTKQMDVSYLLENPHOPASLOATDSOGNTVHALVM 254
 Db 255 ELTKNTNTGSMYGEYPLSFAICMGQHDLFRLMLA---KANLSAQTNTNTALH-LCV 310
 QY 255 ISDNSAENIALVTSMYDGLLOAGARLCPTVOLIEDIRNLQDITPLKLA-KGKIEIFRH 313
 Db 311 IHDK-----MMDLDAVLEAGNI-----RLANKNTALTALAKRKTESIOHL 355
 QY 314 LQRESGLSH--LSRKFEW--CYGPRVSYLYDLASVDCSENSVLEIAFNC----- 362
 Db 356 --ELMDGLIEQILDEK--WKAYG-----RALMRSLSLGFTEFFCCVCAYML 398
 QY 363 -----KSPRRRMV-----LEPLNLKQAKMDLLIPKFLNFCNLIMFTFAVAHQ 412
 Db 399 RPSATPTHLTRGRINDGETESTNSTYLOMHA-----IDTOCHLMY--YSAMPYH 449
 QY 413 PLKKQAAPHKAEVGSMLTGHIILIGIYLLVGOLMYFRRHFWISFIDSYFEI 472
 Db 450 GMFR-----LGCEIMTIYMLFOILLDFGDIRRIGFQKMF-----NFLKAFPAK 493
 QY 473 LFLQALITVVSQVLCPLAI---EMYL-----PLVASVLGMNLNLYTRGFQHTGISV 525
 Db 494 LMFKAFLFIITISIPCRACSFHEFTIDMTMAIISILTYQHFLYMRAPFVGPVVL 553
 QY 526 MIOKVILDLRFLIYLVPLFEGAVALVSLSEAMRPEATPTGNATESVOPMEGOEDEG 585
 Db 554 MYTITATDVRFAMITYSIFLVGSQSYLLFTSCERDS-----TAKKIDPM----- 601
 QY 586 NGAOYRGILE---ASLEFKFTIGMELAFQF---QLHFRGMVLLLLAYVLLTYIL 636
 Db 602 -GSEFNINMENPVDALMTFTIMTIGESVLYREMSACDNFMKMIKLIIFVIFEFVSL 660
 QY 637 LNMILALMSETVNSVADTSMISIKLQKAIISYLEMENYMKCRKKQKRGVNLVGTGRDG 696
 Db 661 QFNLLIAMTTRTYETIFL-TRKEWKQWAOVILMLEMGLSPASRMH---LRLYTRPTG 715

RESULT 5

T37241
 olfactory channel protein osm-9 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 21-Jul-2000
 C:Accession: T37241
 R:Colbert, H.A.; Smith, T.L.; Bargmann, C.I.
 J. Neurosci. 17, 8259-8269, 1997
 A:Title: OSM-9, A novel protein with structural similarity to channels, is required for olfactory behavior in *C. elegans*.
 A:Reference number: 221639; MUID:97477445
 A:Accession: T37241
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-937 <COL>
 A:Cross-references: EMBL:AF031408; NID:92642589; PIDN:AA887064.1; PID:92642590
 C:Genetics:
 A:Gene: osm-9
 A:Map position: IV
 C:Keywords: transmembrane protein

Query Match 9.1%; Score 364; DB 2; Length 937;
 Best Local Similarity 23.9%; Pred. No. 5.4e-20;
 Matches 160; Conservative 101; Mismatches 254; Indels 154; Gaps 23;
 QY 98 LSKTSKLYTDSYTE-----GSTGKTCLMKAVLNLDGVNACTLPLQIDRDSGN 147
 Db 103 LSESVDMQOSRFRKXEHVLMKLNKRGVGEN-LIHLNRPQOVCYETARLL-LKRPFG- 159
 QY 148 PQPLVNAQCTDDYYRGSAHIAIEKRSLOCVKLIVENGANVHARACGRFF----- 198
 Db 160 ---MANDIYLDDEQFGSALHLAIYHDDYETVSLNSKADYVNRACGNFLPDPFKLTN 216
 QY 199 --OKGCTCFYFGELPLSLACTKQMDVSYLLENPHOPASLOATDSOGNTVHALVMIS 256

[illegible]

```

RESULT      6
T24772
hypothetical protein T10B10.7 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
C:Accession: T24772
R:Sims, M.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z19934
A:Accession: T24772
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-519 <MIL>
A:Cross-references: EMBL:Z72514; PIDN:CAA96679.1; CSPDB:GN00028; CESP:T10B10
A:Experimental source: clone T10B10
C:Genetics:
A:Gene: CESP:T10B10.7
A:Map position: X
A:Introns: 47/3; 139/1; 183/1; 215/2; 249/2; 385/3; 440/3

```

Query Match	6.5%	Score 261;	DB 2;	Length 519;
Best Local Similarity	31.2%;	Pred. No. 2.4e-12;		
Matches 90; Conservative	32;	Mismatches 96;	Indels 70;	Gaps 12

QY	157	TDDYVR--GHSALHIAIEKRSLOCVKLLVEGANVWNAACRF--	OKOQG----	204
		:::	:::	
Db	178	SEETYASVGSLPHQAIWNEDLEWVYFLCRGADVHQRCGCSFFC	CADQKSRDLSLEH	237
QY	205	-----CPEGELPLSIAGKQMD-----VSTILENPHQPSLAT	DSQNTVL	249
Db	238	WVDLVQSKTKYGMATGGVPLSPFAACTNQVDCFRLLRMRKADPMP	-----DINGNTVL	291

```

OY 250 HALVIMDSNSENIALYTSMDGLGARGCPYUJEDIRJODLTPLKLARKEGTEI 309
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 292 H-LTVIHD-----LPENMLAVEIGANI-----HVRNNIKLTPPLARLARAKHT 335
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 310 FRHILOREFSGLSHLRKFEWCYGPARYSLYDASVDCSE-----NSVLEITAFHC 362
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 336 YDLLECDMD-----ISWRYGPVYCKAYPLINDYDTINESDGLSNPNSVIANYVGD 386
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
OY 363 KSPRHRRVULLEPLNLQAKMDLLIKPEFLNFCNLIVMFIPTAAVAY 410
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |
Db 367 KVDHIEFFDGL--IEEVLSEKMEFFGKQOLPMSLAGIY---FLAVFY 429
    | : | : | : | : | : | : | : | : | : | : | : | : | : | : | : |

```

RESULT /
T19552
hypothetical protein C29E6.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T19552
R:Dobson, R.
submitted to the EMBL Data Library, May 1996
A:Reference number: Z19141
A:Accession: T19552
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: DNA
A:Residues: 1-1188 <MIL>
A:Cross-references: EMBL:Z72504; PIDN:CAA96603.1; GSPDB:GN00022; CESP:C29E6.2
A:Experimental source: clone C29E6
C:Genetics:
A:Gene: CESP:C29E6.2
A:Map position: 4
A:Introns: 147/2: 473/3: 500/1: 529/1: 584/2: 688/2: 839/3: 975/3: 1132/3

Query Match	5.2%;	Score 209;	DB 2;	Length 1188;
Best Local Similarity	19.1%;	Pred. No. 8.1e-08;		
Matches 152;	Conservative 110;	Mismatches 242;	Indels 290;	Gaps 39

QY	72	NRFPRDR----	LEFNASGVEDBLAGREYLSKSYLTDESEYSGTKCLMKAAVLYNK	128		
Dp	450	NMVRDRONTPIHVIYASNGYLEMMQLQKHGASITQVNEDE-----	TALHRAAIGVA	501		
QY	129	DGVNACILPILQIDRODSGNPRLVNAOCTDYDRKGHALHIAIEKSLQCVLLVNGAN	188			
Dp	502	QLLEWMDIRLELMKDE-----	MGNSALHLAARSGHADATKYLLVNGAD	543		
QY	189	VHAR-----	ACGRFFQKQ-----	OGTCFFGELPILSLAA	217	
Dp	544	KEAANSYOKTRPLOYAVDSGKLETCQRLVANAAGAOIESSDFTKVLTHAAAFVGN-----		596		
QY	218	CTKOMDVVSYLLENPHOPASLOATDSOGNTVLHAIWISDMSKENIA--	LVTSMIDGLL	274		
Dp	597	-----	SLVRFVIA--EGVITDRDEDEGKT--	AFLDACENDHKDVARAELETDOKKNIM	645	
QY	275	QAGARICPTVOLEDIR-----	NLODITPILKLAKEGKIEIFPHILIOREFSGLSHLSRKFTFM	331		
Dp	646	IP-----	CDVIRLDKHNRPVMMKRRTP-----	FRTLL-----	TKFPE- 677	
QY	332	CYGPVAVSLYDLAS--	VDSC-----	EENSYLEITAFH-----	361	
Dp	678	-----	LASFVMDNCIEKSEETSDSTOSVAVNFELDTYMMRCVSDGTEQOL	726		
QY	362	CRKSPH-----		HRRMVVL-----	BLPK-LILOAKWDL	387
Dp	727	GCKSAVDEDFKLEKDAOSVASMTRYKTHPILMLMDAEKHLHLNHPLSAALTKYKMWNL	786			
QY	388	IRPFELNFCNLNLYMELFTAVAYHOPTL-----			KKOAAPHILKAEVGS	430
Dp	787	GRPRYUFAELMYUAVLEFVLSLQYVRRHKAARYVMNNEESYUDEYFEDENECPOJNTKRPD-	845			
QY	431	MULTGHLILILGGLLYLVGQMLFWMRNHYI--	WISFIDSYFELPILFO-ALLTVVSQVL	487		
Dp	846	VVVKIIILQTLACQILVDECFQFQREAFALVMMEMWID-----	CIYSFALLTVVDFSE	899		

QY 488 CF-----IAIEVYLLPLVLSALVGLWLNLLYTRGCHTGYVMIQKVLRLDLLRFLLLYL 543
 Db 900 CSATSGVQNMOMWILALACIFEGWNLILFMIRKMRFGIFVMEVDIV-KTEFFRPFVAV 958
 QY 544 VFLGFAVALVLSQEQAMRPAPT---GPNATESVQPMGEDEGNGAYGILEASLEL 600
 Db 959 LFIITAFSSSEYVILQN--RPEFTSTEMSPKTYVM--MIGE----- 995
 QY 601 FKFTIGMGELEAFOQLHRGM-----VLLLLAVVLLTYILLNLILMSETVNSV 652
 Db 996 FEFT---GIFGDETHAEKKEFGPAHTAVACALFFPFCITMTILMLLVGLAVDDINGV 1052
 QY 653 ATDSMSIWKLOKALS-VLEME-----NGY--WMCRRK 681
 Db 1053 -QEKAEELKRLAMQVDVYLQIBASLHFETQRTKRYATCRATPEPYGKLKGTGFAGWMSNR 1111
 QY 682 QRAGVMTVGTRPD 695
 Db 1112 RRFG--LSVSTDPE 1123

RESULT 8
 JH0588
 Calmodulin-binding protein trpl - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 17-Jul-1992 #sequence_revision 17-Jul-1992 #text_change 24-Sep-1998
 C:Accession: JH0588
 R:Phillips, A.M.; Bull, A.; Kelly, L.E.
 Neuron 8, 631-642, 1992
 A:Title: Identification of a Drosophila gene encoding a calmodulin-binding protein with
 A:Reference number: JH0588; MUID:92232293
 A:Accession: JH0588
 A:Status: nucleic acid sequence not shown
 A:Molecule type: mRNA
 A:Residues: 1-1124 <PHI>
 A:Cross-references: GB:M88185; NID:9469057; PID:9158715
 A:Experimental source: head
 C:Genetics:
 A:Gene: trpl
 A:Cross-references: FlyBase:FBgn0005614
 C:Keywords: calmodulin binding; phosphoprotein; transmembrane protein
 F:341-362/Domain: transmembrane #status predicted <TM1>
 F:374-396/Domain: transmembrane #status predicted <TM2>
 F:462-479/Domain: transmembrane #status predicted <TM3>
 F:512-533/Domain: transmembrane #status predicted <TM4>
 F:549-572/Domain: transmembrane #status predicted <TM5>
 F:643-668/Domain: transmembrane #status predicted <TM6>
 F:710-727/809-823/Region: calmodulin binding #status predicted
 F:722/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 4.9%: Score 197.5; DB 2; Length 1124;
 Best Local Similarity 19.7%: Pred. No. 5.8e-07;
 Matches 167; Conservative 130; Mismatches 281; Indels 249; Gaps 38;
 QY 131 VNACILPLQLDRDSGNPOPL-----VNAQCTDD 159
 Db 25 VGGCCVPL-----GLPQLLEKKFLAVERGDMNVRRILQKALRHQHNINCMRP 77
 QY 160 YRRGSHALHAIERKSQCCKLVLENG-----ANVHARAC-----GRFQKG 201
 Db 78 L--GRRALTLADIDENLEMEVLLVVGVEETKDALHAINAEFVAVELLEHEELIYKEG 135
 QY 202 Q-----GTCEYFGEL-PLSLAAGTKMDVSYLLEN-----PH-----QPRSL 238
 Db 136 EPIYQKVDITAMFAFDITPLMLAHKNNFEIIRILDRGAAPVPHDIRCGCEBECVRL 195
 QY 239 QATDSQNTVLH-----ALVMSDSNAENIALVTSMYDGLQAGARLCTVOLED 288
 Db 196 TAEDSLRHSLSRVNIYALCSPSLICLTNSPSSAPQLSMLNMLALTEBECKSEVYDL 255
 QY 289 IRNLQ-----DLTPKLAKBEGKIEL-----FRHILQREPSGLSHLSRKFTWEC 332

Db 256 RRQCKFAVDLLDQTRTSMELAILLYNDPQMSYEPGRMSLTRLQOAIYKOKFV--A 313
 QY 333 YGPVRSVLS-----YDLASVDSCEBNSVLEIIFHCKSPRHRRHVVLEPINKLL-----QAK 383
 Db 314 HSNIOQLLSIWTYD--GLPGRKRSIVDKVI--CIA-----QVAVLPFLYCLILMCAPICR 365
 QY 384 WDLIPKPEFLNLCN-----LIYMEFTAVA-----YHPTLKQAAPHKAEVGN 429
 Db 366 TGOIMRKPFMFLHASSYLFELFILLVQSQRADDFRIGTGRMKKELAEQELRQGO 425
 QY 430 SMLTGHLILILGCIYLLVQOLMYFRNRHVFITWISFIDSYEILFLQALL--TVSQVL 487
 Db 426 ----TPSKLELIVMYV----IGFVMEVQEIFAVGMSYLRNMNFTDFLRNSLYVSM 477
 QY 488 CELAIEW-----YLP-----LLVSAVLGWLMLLYTTRFQ 518
 Db 478 CLRAFAITQOATEIARDPOMATIPREKWNHDPDQILAEGLRAAANVSALKLVHLESTNP 537
 QY 519 HTGIYSVMIOKVIIRDLRLFLIYLVLEFGRAVALVLSQEQAM-----REAPTPGNAT 572
 Db 538 HGLPQLISLGRNVI-DYKFFFTIVLFAPA--CGINQLMVFPALEKSKCYVLP-- 591
 QY 573 ESVQPMGEDEGNGA-----QYRILEASLELEFTIGMEL-----AFQEQHLFR 619
 Db 592 ----GEADMGSHGDCMKMRFRGNLFESSQSILFMASFGMVGIDFELSGIKSYTRFW 644
 QY 620 GMVILLIAYVLLTYILLNLILMSETVNSVATDSISIKLOKALSIVLEMENYWCRR 679
 Db 645 G--LIMFGSYVINIVTYILLNLILMMSNSYAMIDEHSPTWKFPAN-----TKLMSY 694
 QY 680 KQRAGVMTVGTRPDGSPDERMCFRV-----EEVMAWMEQTLPTLC 722
 Db 695 FEDSA-----TLPPFNVLPSVKWVIRIFRKSSTIDRORSKRRKQEQBFSEYDINMRSLV 750
 QY 723 EDPGAGVPTLENDVLAAPKEDD--DCASEENVVYQVLIQ 762
 Db 751 W-RYVAAMHRRFEN---NPVSEDDINEVKSEINTMYREMLE 787

RESULT 9
 JN0015
 trp protein - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 02-Feb-2001
 C:Accession: JN0015
 R:Wong, F.; Schaefer, E.L.; Roop, B.C.; Lamendola, J.N.; Johnson-Seaton, D.; Shao, D.
 Neuron 3, 81-94, 1989
 A:Title: Proper function of the Drosophila trp gene product during pupal development
 A:Reference number: JN0015; MUID:90148782
 A:Accession: JN0015
 A:Molecule type: mRNA
 A:Residues: 1-1274 <MON>
 C:Comment: This photoreceptor membrane-associated protein is not required for the occ
 C:Genetics:
 A:Gene: trp
 A:Cross-references: FlyBase:FBgn0003861
 A:Map position: 99C5-6
 C:Superfamily: TRPC3 protein
 C:Keywords: glycoprotein; nucleotide binding; P-loop; phosphoprotein; transmembrane P
 F:157-1263/Region: nucleotide-binding motif A (P-loop)
 F:64,70,899/Binding site: carbohydrate (asn) (covalent) #status predicted
 F:191,602,880,883,924/Binding site: phosphate (Ser) (covalent) #status predicted
 F:800,1266/Binding site: phosphate (Thr) (covalent) #status predicted
 Query Match 4.8%: Score 191; DB 2; Length 1274;
 Best Local Similarity 20.2%: Pred. No. 2.2e-06;
 Matches 134; Conservative 99; Mismatches 237; Indels 192; Gaps 30;
 QY 153 NAOCTDYYRHSALHAIERKSQCCKLVLENGANVHARACGRFQKGGCTCYFGLP 212
 Db 62 NINCTDPNMR--SALISAIENENFDLWVILLEHNEY-----GDALLHAISEEY 108

QY 86 GVPEDLAGLPEY-----LSKTSKYITDSEYTBGSGTKCTLMKAV----- 124
 Db 37 GAPPSPGLPSPMAAMAAALYPTDLGASSSSPSSSSSPNEVMALKDQVREKENT 96
 QY 125 LNLKGVNAC-----ILPLLOIDROSGNPQPIVNAOCTDDYRGHSAHIAIEKRSQ 177
 Db 97 LNEKLFPLACKGDIYWKLE--ENSSGD--LNLNCVD--VLGRAAVITTEENLD 149
 QY 178 CVKLLVENGAVNARACGRFQKGGTGF--YFGEPLSLAACKOMDVSYLLE----- 230
 Db 150 ILQOLLIDYG-----CQKLMERIONPEYSTTMQVAPYIILAHNNNYELLMLKQDVS 202
 QY 231 -NPH-----OPASLOATDSOGNTVLAHVMISD 257
 Db 203 PKPAVCECTLSAKKKKDSLRSRFLDIYRCLASPALIMLTEE--DPILRAFEISAD 260
 QY 258 -----NSAENIALVTSMYDGLLOAGARLCPVOEDIRN----- 291
 Db 261 LKELSLVEFRNDYEEELARCKMFAKDLAQR--NSRELEVILNMTSSDEPLDKGL 318
 QY 292 --LQDLPLKLAKEGKIEIFRHILQREFSGLSHSRKTEMCYGPVRSYDLASVDS 349
 Db 319 EERNLISRLKAIKYN-----QKEF-----VSQSN 344
 QY 350 EENSVELIAFHCKSPHRHR-----MYLE-----PLNKL-----QAKDILLPKFEL 393
 Db 345 QO--FLMTVMFGOMSGYRKPKTKIMTVLTVGIFWPVLSICYLIAPKSGRIIHPFM 402
 QY 394 NELCNLIYMEFT-----ANAYHQPILKQAARHLKAEVNSMLTGILLGCIYL 447
 Db 403 KFTIHGASYFFLLILLNLISLVNEDK-----KNTMGPALERIDYLL-----WT 448
 QY 448 VQGLMFWRRHVFIMISFIDSYFE-----ILFLQA--LLTVSQQVLCFLAI-----E 493
 Db 449 IGMIMSDIKR--LMYGLIEDFLESRNQSLFVNNSLYLAFLAKVVAHNKHFADK 505
 QY 494 W--LPLLSALVIGWMLLYTTRGFOHTGIYSW-----IQKYLRLDLPLLIYVEL 546
 Db 506 WDAHPTLVAGELFAFANVLSYLRLEFMYTTSILGPLQISMGMLODFGFLCMFLLYL 565
 QY 547 FGFVALVLSQGEAMREPAPGPNATESVQPMGEODEGNAOYRGILEASLELFKFTIG 606
 Db 566 FSFITGLTOLYDKGT-----TSKQKDCVGFCEQOSNDT---FHSFGTCFALFWYFS 617
 QY 607 MGLAFL-----QEOLHFRGAVLLLLAYVLLTYILLMLLALMSETVNSVATDSMS 658
 Db 618 LAHVAIFVTRFSYGBELQSFYGAIV--LYGTNVVVIVITRLVLAAMLHKSQOLLANHEDK 675
 QY 659 IWLQKA 665
 Db 676 EMKFARA 682
 RESULT 12
 H71274
 Probable ankryrin - syphilis spirochete
 C:Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 10-Sep-1999
 C:Accession: H71274
 R:Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwin
 rson, J.; Khalak, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Uterback, T.; McD
 they, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
 Science 281, 375-388, 1998
 A:Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
 A:Reference number: A71250; MUID:9832770
 A:Accession: H71274
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-934 <COL>
 A:Cross-references: GB:AE001254; GB:AE000520; NID:g323148; PIDN:AAC65803.1; PID:g3232314
 A:Experimental source: strain Nichols
 C:Genetics:
 A:Gene: TP0835

C:Superfamily: syphilis spirochete probable ankryrin; ankryrin repeat homology
 F:281-313/Domain: ankryrin repeat homology <AN22>
 Query Match 4.28; Score 167.5; DB 1; Length 934;
 Best Local Similarity 27.68; Pred. No. 9,7e-05;
 Matches 72; Conservative 37; Mismatches 95; Indels 57; Gaps 10;
 QY 76 RDRLFNNAVSRCVPEDLAGLPEYLSKTSKYITDSEYTBGSGTKCTLMKAVL----- 126
 Db 448 RDLPHAVASRAHNDIV---KFLFRERQM---IARPMNTILHAYVANDDAVGEFL 500
 QY 127 LKGVNAC-----ILPLLOIDROSGNPQPL-----VNAOCTDDYRGHSAHIAIEKR 174
 Db 501 MREGADIFSTVNHVSPKLTALTSGRGDWILTAANYHAODTG--GNPRLHACEWK 556
 QY 175 SLOCVKLLVENGAVNARACGRFQKGGTGFYFGEPLSLAACKOMDVSYLLE-NPH 233
 Db 557 LQAINGILRRGAETEARNLQ-----ETPLFSAVSDAEVISILHPQAG 603
 QY 234 QPASLOATDSOGNTVLAHVMISDPSAENIALVTSMYDGLLOAGARLCPVOEDIRNLQ 293
 Db 604 NPALVDARDAGNTVLAHCYRMSALRSADVL-----IREADAR---HVSILNARL 652
 QY 294 DLPLKLAKEGKIEIFRHIL 314
 Db 653 GKPLHLAARAGNVDFIRLL 673
 RESULT 13
 T23729
 Hypothetical protein M05B5.6 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999
 C:Accession: T23729
 R:Gardner, A.
 Submitted to the EMBL Data Library, April 1996
 A:Reference number: 219790
 A:Accession: T23729
 A:Molecule type: DNA
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Residues: 1-481 <ML>
 A:Cross-references: EMBL:Z11265; PIDN:CAA95836.1; GSPDB:GN00019; CESP:M05B5.6
 A:Experimental source: clone M05B5
 C:Genetics:
 A:Gene: CESP:M05B5.6
 A:Map position: 1
 A:Introns: 50/2; 99/1; 133/3; 185/3; 229/3; 402/3; 457/3
 Query Match 3.9%; Score 157.5; DB 2; Length 481;
 Best Local Similarity 23.08; Pred. No. 0.00024;
 Matches 90; Conservative 58; Mismatches 149; Indels 95; Gaps 17;
 QY 291 NLQDLTPKLAKEGKIEIFRHILQREFSGLSHSRKTEMCYGPVRSYDLASVDSCE 350
 Db 57 NMRKLRLKSGKKWG---VIRH-----PYILNVYVQKLIDCAFFYS-- 94
 QY 351 ENSVLEIIFAFCKSPHRHNVLEPINKLQAKDILLPKFELNLCILYMFETNAVAY 410
 Db 95 ---LHILAF-----LVFLLAHMHVSRLNEKDL---TTFGTGIF 130
 QY 411 HQPLKQAARHLKAEVNSMLTGILLGCIYLLVQGLMFWRRHVFIMISFIDSYF 470
 Db 131 MFLVLKGF---IKARITKS--VSTWFIYAFPCFNIFTWATLAVWLPVFE---GYDDVHL 182
 QY 471 EILFLQALLTVSQQVLCFLAIEWYLLPLLSALVIGWMLLYTTRGFOHTGIYSVMIOKV 530
 Db 183 EVKRI-----VTWFLPII--AIIAMWMLLIYMRK--SPFGIYFMMYR- 222
 QY 531 ILRDLRLFLLIYLVLFGEFAVALVLSQGEAMREPAPGPNATES---VQPMGEODEGNG 587
 Db 223 ILRSFAHIAITWIPLLAFSFAFLIMRDGVCKMPPLIDQOTEMTMVQTM----- 273

THIS PAGE BLANK (USPTO)